УДК 519.7

Т.Ф. Басканова[1], *Ю.П. Ланкин*[2]
[1]Сибирский федеральный университет, г. Красноярск, Россия
tfbask@mail.ru
[2]Институт биофизики СО РАН, г. Красноярск, Россия
lan7@mail.ru

Нейросетевой анализ непрерывных потоков нестационарных данных

В статье предложены структуры динамических (рекуррентных) и статических (слоистых) нейронных сетей для обработки стационарных и нестационарных сигналов. Эти структуры позволяют работать на больших временных интервалах и в непрерывном режиме, аппроксимировать временные ряды и функциональные зависимости, а также выполнять распределенную обработку входных данных.

Введение

Искусственный интеллект возник как способ имитации естественного интеллекта, базирующегося на нейронных сетях человеческого мозга и ставшего основой человеческой культуры, науки и искусства. Нейрокибернетика, возникшая во второй половине сороковых годов прошлого века, дала начало исследованиям по технической имитации решения интеллектуальных задач человеком. В шестидесятые годы эта волна исследований уступила место имитации элементов человеческого мышления на основе формальной логики. В начале восьмидесятых (после работ Дж. Хопфилда), в связи с появлением алгоритмов, которые позволили обучать «скрытые» (не связанные со входами и выходами нейросети) слои нейронов, возникла вторая волна интереса к нейронным сетям. Эти алгоритмы значительно расширили возможности нейронных сетей, породив нейроинформатику – науку об обработке информации и управлении с помощью нейронных сетей. В дальнейшем нейросети стали активно внедряться в системы искусственного интеллекта с целью расширения возможностей формально-логических методов при решении трудно формализуемых задач [1]. В программе «Real World Computing (RWC)»/«Вычисления в реальном мире» для шестого поколения компьютеров «огромную роль (до 30 – 40 процентов ее содержания) отводят исследованию естественных и созданию искусственных нейронных сетей».

В современной нейроинформатике существует множество алгоритмов для решения широкого круга научно-технических задач. Спектр приложений стремительно расширяется и активно вторгается в медицину, экологию, биологию и даже в искусство. Доказан ряд теорем, обосновывающих широкие возможности нейронных сетей [2-4], и пишутся толстые учебники [5]. Показано, что тенденции вычислительной техники фокусируются на нейрокомпьютерных технологиях [6]. Предложен способ описания иерархических архитектур нейронных сетей [7] и поисковый алгоритм для их обучения [8] (направлена в печать детерминированная версия алгоритма). Существует амбициозный проект, претендующий на построение полномасштабной модели человеческого мозга [9]. Возникло обратное движение от искусственных нейронных сетей

к их биологическим прототипам. В медицину активно внедряются нейросетевые экспертные системы [10], а в психологии имеются попытки использования аналогий с искусственными нейронными сетями для решения психологических задач [11], [12].

Вместе с тем имеются определенные трудности, преодоление которых требует больших усилий и не всегда дает ожидаемые результаты. В частности, с рядом проблем сопряжено решение задач, связанных с обработкой нестационарных сигналов и пространственно-временных паттернов большой длительности, включая режимы непрерывного функционирования, присущие человеческому мозгу, который породил естественный интеллект.

Состояние области исследований: нейронные сети и спектральный анализ

Возможность представления *нейронных сетей* в виде рядов, как будет показано ниже, естественным образом подталкивает к сопоставлению с традиционными методами *спектрального анализа*. В частности, гармонического анализа / синтеза с помощью рядов Фурье. Здесь допустимы не только сопоставления, но и прямое перенесение подобных представлений в нейроалгоритмы, как это сделано в известной программе Модель [13], нейроны которой описываются уравнениями:

$$Y_i = \sum_{j=1}^{N} a_j \sin\left(\sum_{k=1}^{M} b_{kj} X_k\right),$$

где Y_i – выход i-го нейрона; N – число нейронов; M – размерность входного вектора; X_k – входной вектор; a, b – подстроечные коэффициенты нейронной сети.

В работе [14] показано, что существуют системные инварианты, определяющие общие правила построения быстрых алгоритмов (Фурье, Уолша – Адамара, Хаара, Виленкина – Кристенсона и др.), и уточняется, что с позиций сегодняшнего дня быстрое перестраиваемое преобразование можно рассматривать как специфичную, многослойную нейронную сеть с прореженными связями и линейными функциями активации. При трансформации быстрого перестраиваемого преобразования в нейронную сеть в каждый слой добавляются нелинейные функции активации. Исследование инвариантных свойств быстрых алгоритмов также выполнено на примере быстрого преобразования Фурье (БПФ).

Преобразование Фурье получило широкое распространение в науке и технике, особенно электронике [15], теории управления [16], цифровой обработке сигналов [17], [18] и изображений [19], цифровой голографии [20], физике [21], [22] и др. Оконное преобразование Фурье использовалось для распознавания речи [23], исследования свойств зрения [24] и т.д.

Вместе с тем, при всей своей популярности, преобразование Фурье имеет ряд недостатков, которые особенно заметно проявляются на нынешнем этапе научно-технического развития. Дело в том, что современные системы обработки информации и управления все в большей степени ориентируются на работу со сложными сигналами в реальном времени в условиях помех. Повышенные требования к точности и быстродействию при обработке нестационарных сигналов (а таковых сейчас большинство) создают серьезные трудности для применения БПФ [25], включая и его «оконные» версии.

Среди проблем, возникающих при использовании БПФ для обработки сложных сигналов, выделим следующие [25]:

– базисной функцией для разложения в ряд Фурье является гармоническое (синусоидальное) колебание, которое математически определено в интервале времени от $-\infty$ до $+\infty$ и имеет неизменные во времени параметры;

– в результате, отдельные особенности сигнала (например, разрывы или пики) вызывают незначительное изменение частотного образа сигнала во всем интервале частот от $-\infty$ до $+\infty$, которые «размазываются» по всей частотной оси, что делает их обнаружение по спектру практически невозможным;

– такая плавная базисная функция, как синусоида, в принципе не может представить перепады сигналов с бесконечной крутизной, хотя такие сигналы (например, прямоугольные импульсы) применяются очень широко;

– единственным приспособлением к представлению быстрых изменений сигналов, таких как пики или перепады, является резкое увеличение числа гармоник, которые оказывают влияние на форму сигнала и за пределами локальных особенностей сигнала;

– по составу высших составляющих спектра практически невозможно оценить местоположение особенностей на временной зависимости сигнала и их характер;

– для нестационарных сигналов (а таковых сейчас большинство) трудности многократно возрастают;

– в случае оконного преобразования Фурье действует принцип неопределенности, согласно которому улучшение разрешения по одному параметру (частоте или времени) ведет к ухудшению разрешения по другому параметру [26].

Очевидно, что некоторые проблемы Фурье-преобразования характерны и для нейронных сетей традиционной архитектуры, сопоставление с которыми, приведенное в работе [14], обсуждалось выше.

Одним из эффективных способов решения задачи анализа нестационарных сигналов является использование вейвлет-анализа, позволяющего отражать локальные особенности сигналов и решающего другие описанные проблемы.

Для нас, в данном случае, представляет интерес не попытка прямого переноса математического аппарата вейвлетов [27] на нейронные сети, что представляется неразумным, а трансляция некоторых базовых идей вейвлет-анализа на нейросетевой базис.

Цель данной работы – предложить формальное описание динамических (рекуррентных) и статических (слоистых) нейронных сетей для обработки стационарных и нестационарных сигналов. Эти структуры позволяют эффективно работать на больших временных интервалах и в режиме непрерывного функционирования.

Постановка задачи. Анализ современного состояния области исследований позволил провести определенные параллели между методами спектрального анализа и современными нейронными сетями. Несмотря на большие возможности, алгоритмы современной нейроинформатики содержат определенные ограничения, не позволяющие полностью раскрыть потенциал нейросетевых методов в обработке и анализе информации, а также управлении сложными системами. Среди них – ограничения по формированию сложных сигналов при непрерывном функционировании и трудности с обработкой нестационарных сигналов, сходные с проблемами традиционного спектрального анализа.

Упомянутые трудности ограничивают дальнейший рост возможностей нейронных сетей в решении сложных задач, таких как распознавания речи и визуальных

паттернов в условиях динамичных изменений окружающей обстановки, управление сложной динамикой распределенных объектов и др.

Для преодоления указанных трудностей ставится задача разработки математических представлений, позволяющих предложить формальное описание нейронных сетей для работы на больших временных интервалах и в режимах непрерывного функционирования, распределенной обработки информации и обработки нестационарных сигналов.

Динамические сети для больших интервалов времени

Одним из существенных ограничений большинства современных нейросетевых алгоритмов является неспособность, либо ограниченные возможности обучения и функционирования на больших интервалах времени. Между тем, дальнейший прогресс нейроинформатики с необходимостью обусловлен наличием такой возможности, позволяющей нейронным сетям эффективно решать задачи моделирования и анализа сложных сигналов реального мира, адаптивного управления поведением сложных объектов и др.

Для построения нейросетей, способных работать со стационарными сигналами на больших интервалах времени, рассмотрим модель, функционирующую в дискретном времени. Простейший пример такой нейронной сети приведен на рис. 1. Пусть функционирование i-го формального нейрона нейросети описывается уравнением:

$$\alpha_i^k = arctg(\gamma_i \rho_i^k + \omega_i \rho_i^{k-1}), \qquad (1)$$

где γ_i и ω_i – подстроечные коэффициенты, а

$$\rho_i^k = \sum_j \alpha_j^{k-1} x_{ij} + A_i^k, \qquad (2)$$

где α_j – сигналы от других нейронов; k – дискретный момент времени функционирования сети; x_{ij} – вес связи от i-го нейрона к j-му; A_i – внешний входной сигнал i-го нейрона.

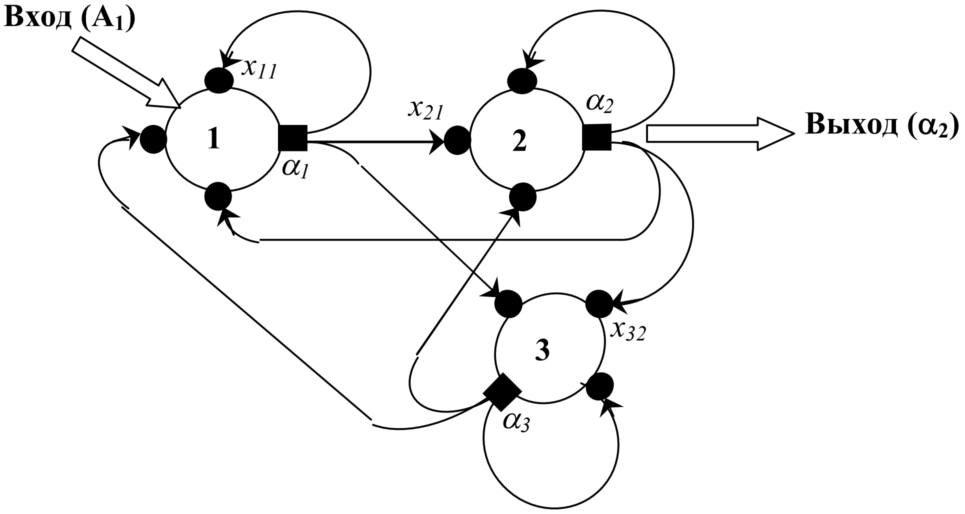


Рисунок 1 – Пример простейшей полносвязной нейронной сети из трех нейронов [28]: 1 – входной нейрон; 2 – выходной нейрон; 3 – ассоциативный нейрон, который не связан непосредственно с входом и выходом

Преобразуем уравнение (1) к следующему виду:

$$tg\alpha_i^k - \gamma_i A_i^k - \omega_i A_i^{k-1} = \gamma_i \sum_j \alpha_j^{k-1} x_{ij} + \omega_i \sum_j \alpha_j^{k-2} x_{ij}, \quad (3)$$

Просуммировав по i уравнения (3), получим:

$$\sum_i \left(tg\alpha_i^k - \gamma_i A_i^k - \omega_i A_i^{k-1} \right) - \sum_i \sum_j \left(\gamma_i \alpha_j^{k-1} + \omega_i \alpha_j^{k-2} \right) x_{ij} = 0. \quad (4)$$

Уравнение (4) демонстрирует представление нейронной сети в виде ряда, с помощью которого можно аппроксимировать любую непрерывную функцию на больших интервалах времени. Напомним, что такое представление эффективно для стационарных сигналов.

Динамические сети для нестационарных сигналов

Как уже указывалось в предыдущем параграфе, уравнение функционирования (1) эффективно для представления стационарных сигналов. Однако на практике часто встречаются сигналы с изменяющимися во времени свойствами. Для качественной аппроксимации таких сигналов нейронной сетью, по аналогии с вейвлет-преобразованием, уравнение функционирования нейрона может быть дополнено членом ψ_i:

$$\alpha_i^k = \psi_i(k)\, arctg(\gamma_i \rho_i^k + \omega_i \rho_i^{k-1}), \quad (5)$$

где ψ_i является аналогом базисных функций при вейвлет-преобразовании, а оставшаяся часть выражения – эквивалентом вычисления коэффициентов при функции ψ_i, содержащих информацию о конкретном сигнале. Одним из возможных способов представления функции ψ_i является следующий:

$$\psi_i(k) = \mu_i\,(1 - (\nu_i k - \phi_i)^2)\, e^{-\frac{(\nu_i k - \phi_i)^2}{2}}, \quad (6)$$

где μ_i – коэффициент масштабирования; ν_i – коэффициент сжатия / растяжения; φ_i – коэффициент смещения по оси дискретного времени.

В данном случае ряд, представляющий нейронную сеть, примет следующий вид:

$$\sum_i \left(tg\frac{\alpha_i^k}{\psi_i(k)} - \gamma_i A_i^k - \omega_i A_i^{k-1} \right) - \sum_i \sum_j \left(\gamma_i \alpha_j^{k-1} + \omega_i \alpha_j^{k-2} \right) x_{ij} = 0. \quad (7)$$

Одним из отличий предложенного подхода от традиционного вейвлет-анализа является отсутствие жесткой необходимости в конструировании базисных функций, поскольку коэффициенты при уравнении (6) могут быть получены с помощью адаптивной процедуры наравне с подстроечными параметрами x_{ij}.

Статические сети для нестационарных сигналов

Уравнение (5) легко модифицировать для статической (нерекуррентной) сети с целью аппроксимации в дискретном времени заданной функциональной зависимости или сигнала:

$$\alpha_i^k = \psi_i(k)\, arctg(\rho_i^k). \quad (8)$$

При этом уравнение (6) остается неизменным. Здесь, для упрощения, при использовании слоистой сети предполагается работа с одним скрытым слоем $n = 2$. На уравнение (2) вводится условие: $i \neq j$.

Статические сети для распределенной обработки

Интересные результаты могут быть получены модификацией предложенных представлений на случай задач, не содержащих времени. Так, вместо времени может быть введена зависимость от номера класса входных данных. Очевидно, что в этом случае входные данные нейросети перед подачей на нее должны классифицироваться, например, с помощью алгоритма [29]. Перепишем уравнение (5) следующим образом:

$$\alpha_i^k = \theta_i(n)\, arctg(\rho_i^k)\,, \tag{9}$$

где n – номер класса для входных данных. Уравнение (6) при этом, примет вид:

$$\theta_i(n) = \mu_i\,(1-(\nu_i n-\phi_i)^2)\, e^{-\frac{(\nu_i n-\phi_i)^2}{2}}\,. \tag{10}$$

Здесь, как и в предыдущем параграфе, предполагается работа с одним скрытым слоем $n = 2$ и на уравнение (2) вводится условие: $i \neq j$.

Такая нейросеть будет избирательно реагировать на входные данные, распределяя их обработку по пространству сети.

Заключение

Использование предложенных математических методов и типов формальных нейронов дает возможность строить архитектуры нейронных сетей, которые во многих случаях позволяют решать задачи повышенной сложности, не прибегая к громоздким комбинациям разнородных методов, усложняющих и удорожающих систему, увеличивающих время разработки, а также сказывающихся на ее надежности. Особенно это касается иерархических архитектур (для нейросетей такое описание предложено в [7]), необходимых при решении задач повышенной сложности.

Предложенные решения позволяют нейросетям эффективно работать на больших интервалах времени и в непрерывных режимах, вести распределенную обработку входных информационных потоков и работать с нестационарными входными сигналами, характерными для реальной окружающей среды, в которой зачастую приходится работать системам искусственного интеллекта. Такими задачами являются, например, задачи распознавания речи и визуальных паттернов в условиях динамичных изменений окружающей обстановки.

Литература

1. Дунин-Барковский В.Л. Нейрокибернетика, нейроинформатика, нейрокомпьютеры / Дунин-Барковский В.Л. // В кн. Нейроинформатика. – Новосибирск : Наука. Сибирское предприятие РАН, 1998. – С. 5-17.
2. Горбань А.Н. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей / А.Н. Горбань // Сибирский журнал вычислительной математики. – Новосибирск : РАН. Сиб. отделение, 1998. – № 1. – С. 11-24.
3. Колмогоров А.Н. О представлении непрерывных функций нескольких переменных суперпозициями непрерывных функций меньшего числа переменных / Колмогоров А.Н. // Докл. АН СССР. – 1956. – Т. 108, № 2. – С. 179-182.
4. Маккаллок У.С. Логическое исчисление идей, относящихся к нервной активности / У.С. Маккаллок, У. Питс // Автоматы. – М. : Изд-во иностр. лит-ры, 1956. – С. 362-384.
5. Хайкин С. Нейронные сети: полный курс / Хайкин С. ; пер с англ. – М. : Издат. дом «Вильямс», 2006. – 1104 с.
6. Барцев С.И. Принцип двойственности в организации адаптивных сетей обработки информации / Барцев С.И., Гилев С.Е., Охонин В.А. // Динамика химических и биологических систем. – Новосибирск : Наука. Сиб. отд-ние, 1989. – С. 6-55.

7. Басканова Т.Ф. Иерархические нейронные сети, как средство решения трудноформализуемых задач искусственного интеллекта / Т.Ф. Басканова, Ю.П. Ланкин, С.В. Комиссаров // Искусственный интеллект. – 2009. – № 1. – С. 100-111.
8. Lankin J.P. Algorithms of self-adaptation for atmospheric model designing / J.P. Lankin, T.F. Baskanova // SPIE. – 2004. – Vol. 5397. – P. 260-270.
9. Artificial Development (AD). – Режим доступа : http://www.ad.com.
10. Горбань А.Н. Нейронные сети на персональном компьютере / А.Н. Горбань, Д.А. Россиев. – Новосибирск : Наука. Сибирская издательская фирма РАН, 1996. – 276 с.
11. O'Конор Джозеф. Нейронные сети и стратегии НЛП / O'Конор Джозеф., Ван дер Хорст Брайен // «NLP» нейро-лингвистическое программирование. Вестник современной практической психологии. – Заказ № 551. – 1998. – С. 27-39.
12. Дилтс Роберт. Моделирование с помощью НЛП / Дилтс Роберт; пер. с англ. – СПб. : Питер, 2001. – 288 с. – (Серия «Практикум по психотерапии»).
13. Okhonin V. Neural network based approach to the Evaluation of Degradation Lifetime / V. Okhonin, S. Okhonin, A. Ils [and all.] // Neural Network World. – 2001. – Vol. 11, № 2. – P. 145-151.
14. Дорогов А.Ю. Структурные и топологические инварианты быстрых перестраиваемых преобразований / А.Ю. Дорогов // VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006». Научная сессия МИФИ-2006: в 3-х частях. Ч. 1. – М. : МИФИ, 2006. – С. 39-50.
15. Манаев Е.И. Основы радиоэлектроники / Манаев Е.И. – М. : Советское радио, 1976. – 480 с.
16. Никулин Е.А. Основы теории автоматического управления. Частотные методы анализа и синтеза систем / Никулин Е.А. – СПб. : БХВ-Петербург, 2004. – 640 с.
17. Блейхут Р. Быстрые алгоритмы цифровой обработки сигналов / Блейхут Р.; пер. с англ. – М. : Мир, 1989. – 448 с.
18. Даджион Д. Цифровая обработка многомерных сигналов / Д. Даджион, Р. Мерсеро; пер. с англ. – М. : Мир, 1988. – 488 с.
19. Рудаков П.И. Обработка сигналов и изображений. MATLAB 5.x / П.И. Рудаков, В.И. Сафонов / под общ. ред. В.Г. Потемкина. – М. : ДИАЛОГ-МИФИ, 2000. – 416 с.
20. Ярославский Л.П. Цифровая голография / Л.П. Ярославский, Н.С. Мерзляков. – М. : Наука, 1982. – 219 с.
21. Савельев И.В. Основы теоретической физики. Механика, Электродинамика / Савельев И.В. – М. : Наука. Гл. ред. физ.-мат. лит., 1975. – Т. 1. – 416 с.
22. Вихман Э. Квантовая физика / Вихман Э.; пер с англ. – М. : Наука. Гл. ред. физ.-мат. лит., 1974. – 415 с.
23. Kohonen T. The «Neural» Phonetic Typewriter / Kohonen T. // IEEE Computer, March 1988. – P. 11-22.
24. Глезер В.Д. Зрение и мышление / Глезер В.Д. – СПб. : Наука, 1993. – Изд. 2-е. – 284 с.
25. Дьяконов В. MATLAB. Обработка сигналов и изображений. Специальный справочник / В. Дьяконов, И. Абраменкова. – СПб. : Питер, 2002. – 608 с.
26. Дьяконов В. Вейвлеты. От теории к практике / Дьяконов В. – М. : САЛОН-Пресс, 2004. – 400 с.
27. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB / Смоленцев Н.К. – М. : ДМК Пресс, 2005. – 304 с.
28. Барцев С.И. Теоретические модели в решении ключевых задач биосферики : дисс. на соиск. уч. степ. д.ф.-м.н. / Барцев С.И. – Красноярск : Институт биофизики СО РАН, 2005. – 221 с.
29. Ланкин Ю.П. Обучение без учителя на основе нейронных сетей с самостоятельной адаптацией / Ю.П. Ланкин, Т.Ф. Басканова // Реляторные, непрерывно-логические нейронные сети и модели. Труды международной конференции «Континуальные алгебраические логики, исчисления и нейроматематика в науке, технике и экономике – КЛИН-2002». – Ульяновск : УлГТУ, 2002. – Том 3. – С. 39-41.

Т.Ф. Басканова, Ю.П. Ланкін

Нейромережний аналіз безперервних потоків нестаціонарних даних

У статті запропоновані структури динамічних (рекурентних) та статичних (шаруватих) нейроних мереж для обробки стаціонарних та нестаціонарних сигналів. Ці структури дозволяють працювати на великих часових інтервалах та у безперервному режимі, апроксимувати часові ряди та функціональні залежності, а також виконувати розподілену обробку вхідних даних.

Статья поступила в редакцию 14.07.2009.